Structure and dynamics of a designed helix-loop-helix dimer in dilute aqueous trifluoroethanol solution. A strategy for NMR spectroscopic structure determination of molten globules in the rational design of native-like proteins

Susanne Olofsson and Lars Baltzer

Background: The overwhelming majority of engineered amino acid sequences designed to fold into well defined tertiary structures show the hallmarks of molten globules. Although imperfectly folded, the structures of these polypeptides are of considerable interest in assessing the predictive power of design strategies and in understanding the structural basis for the formation of proteins with native-like properties. This paper describes a strategy for the structural characterization of molten globules by NMR spectroscopy applied to the study of SA-42, a polypeptide with 42 amino acids that folds into a hairpin helix-loop-helix dimer.

Results: The ^{1}H NMR spectrum of SA-42 was assigned in several mixtures of water and trifluoroethanol (TFE) (0–30 vol%) and small amounts of TFE were shown to have a significant effect on the spectrum. The secondary and supersecondary structures of SA-42 were determined. In aqueous solution a helix-loop-helix dimer is formed, but in 30 vol% of TFE the populations of hairpin dimers are negligible and SA-42 is monomeric, folding into two non-interacting helical segments. In solutions containing less than 3 vol% of TFE the structure is very similar to that in water and the structural information may be used to develop the motif in aqueous solution. Less well ordered amino acid residue sidechains in the hydrophobic core were identified. Helix distortion in the tetrahelix bundle was found to be small.

Conclusions: Detailed information about molten globule structures in aqueous solution can be obtained from NMR spectroscopy if the spectra are assigned in dilute TFE solution. On the basis of the NMR spectroscopic analysis, the solution structure of SA-42 was found to be close to the designed one. A route for developing native-like properties in SA-42 is suggested based on the identification by NMR spectroscopy of some less well ordered amino acid sidechains in the hydrophobic core and on the observed structural rigidity of the two helices.

Address: Department of Organic Chemistry, Göteborg University, 412 96 Göteborg, Sweden.

Correspondence: Lars Baltzer
e-mail: baltzer@oc.chalmers.se

Key words: Helix-loop-helix dimer, molten globule, NMR spectroscopy, rational design

Received: 28 Mar 1996
Revisions requested: 22 Apr 1996
Revisions received: 09 Jul 1996
Accepted: 10 Jul 1996

Published: 23 Aug 1996
Electronic identifier: 1359-0278-001-00347

Folding & Design 23 Aug 1996, 1:347–356

Introduction

The *de novo* design of polypeptides and proteins [1,2] is an important goal because it tests our understanding of the folding problem and because it has great potential in, for example, the engineering of catalysts [3,4], immunogenic determinants [5], membrane ion channels [6,7] and DNA-binding proteins [8]. To date, several designed polypeptides with supersecondary structures have been reported [1,4–11], but the design of polypeptides with protein-like properties remains a considerable challenge. It was only recently that the first designed four-helix bundle with characteristics of a native protein, α_2D, was reported [9]. The overwhelming majority of designed polypeptides show the dynamic properties of molten globules. They have poorly dispersed NMR spectra with broadened resonances, poorly defined melting points, they bind 8-anilino-1-naphthalene-sulphonic acid (ANS) etc. A molten globule is characterized by a large amount of secondary structure but a poorly ordered hydrophobic core. In order to develop our understanding of how to design native-like proteins from imperfectly folded polypeptides, poorly ordered sidechains must be identified so that the packing of the hydrophobic interfaces can be fine-tuned in a rational way.

NMR spectroscopy is uniquely suited to provide structural information in solution but it has not been widely used in the study of designed proteins. The effect of dynamics on the NMR spectrum of a folded polypeptide is to broaden the resonances and cause overlap and to make magnetization transfer inefficient. Multidimensional

NMR techniques will not, therefore, provide the resolution and sensitivity necessary. To overcome these problems in the study of SA-42, a polypeptide with 42 amino acids that folds into a hairpin helix-loop-helix dimer in aqueous solution [10], small amounts of trifluoroethanol (TFE) were added to aqueous solutions of the polypeptide. TFE is a solvent that reduces amide proton exchange rates [12] and as a result the line broadening of amide protons is reduced, the resolution is improved and the intensity is increased. The physical reasons for the reduced exchange rates are not understood but may be due to the difference in hydrogen-bonding properties of TFE mixtures relative to those of water. TFE is also known to increase the helical content of sequences with helix propensity [13] and to disrupt the tertiary structure of proteins [14]. Here, we report that the addition of small amounts of TFE has significant effects on the ^{1}H NMR spectrum of SA-42 and little effect on its structure. The use of TFE titrations may, therefore, be of general use in NMR spectroscopic studies of molten globule structures. The amide region of the spectrum is improved due to reduced amide proton exchange rates as indicated above. The resolution is also improved in the αH region, probably due to the effect on the helical structure by the solvent. The protons in the sidechains of the hydrophobic residues are better resolved since the solvent drives the equilibrium towards conformers with more solvent-exposed sidechains. In this case, the increased resolution is due to a more rapid sampling of conformational space.

The assignment of the ^{1}H NMR spectrum of SA-42 has been carried out in several mixtures of water and TFE (0–30 vol% TFE) and the secondary and supersecondary structures have been determined. In aqueous solution, SA-42 folds into a hairpin helix-loop-helix motif that dimerizes. In solutions with more than 30 vol% TFE, the supersecondary structure of SA-42 is lost and it forms monomers with two non-interacting helices connected by a loop. The helical segments have been identified and the helix distortion in the tetrahelix bundle has been shown to be small. Sidechains that are poorly ordered in the hydrophobic core have been identified and a route for the development of native-like properties in SA-42 is suggested.

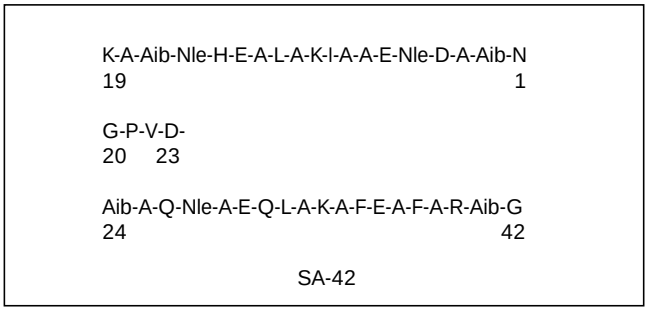
SA-42 and two designed helix-loop-helix dimers developed from SA-42 by exchange of three and five amino acid residues were recently shown to exhibit catalytic activity in 10 vol% TFE in 90% 50 mM aqueous Bis-Tris buffer at pH 5.8 and 290 K [4].

Results

Design of SA-42

The design of SA-42 has been described in detail previously [10]. In short, it is made up of two amphiphilic helices connected by a short loop. The helices were designed according to established principles [1]. Interhelical hydrophobic interactions were obtained through the incorporation of one isoleucine, two leucines, three norleucines and two phenylalanines in the sequence (Fig. 1). The amino acid composition was varied as much as possible to simplify the assignment of the ^{1}H NMR spectrum and the only two phenylalanines in the sequence were placed at the lower end of helix II to induce shift dispersion in the folded state (Fig. 2).

Figure 1

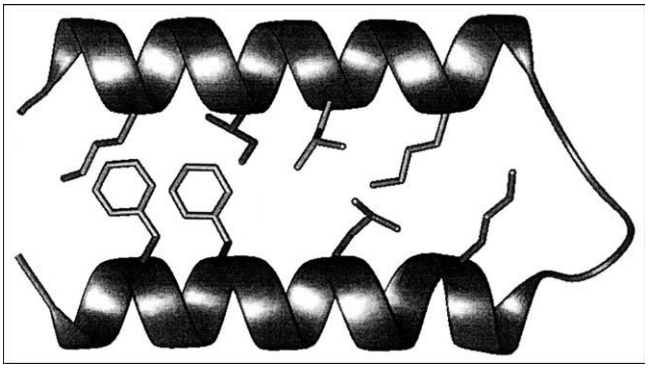


Amino acid sequence of SA-42 using the one-letter amino acid code. Aib is α-aminoisobutyric acid and Nle is norleucine.

The 1D ^{1}H NMR spectrum and the dynamics of SA-42

The ^{1}H NMR spectrum of SA-42 was recorded as a function of the concentration of TFE (Fig. 3). The addition of TFE reduced the linewidth but also the chemical shift dispersion in the ^{1}H NMR spectrum of SA-42. The net effect was to increase the spectral resolution. A similar effect was obtained upon raising the temperature, but higher temperature also led to an increased rate of

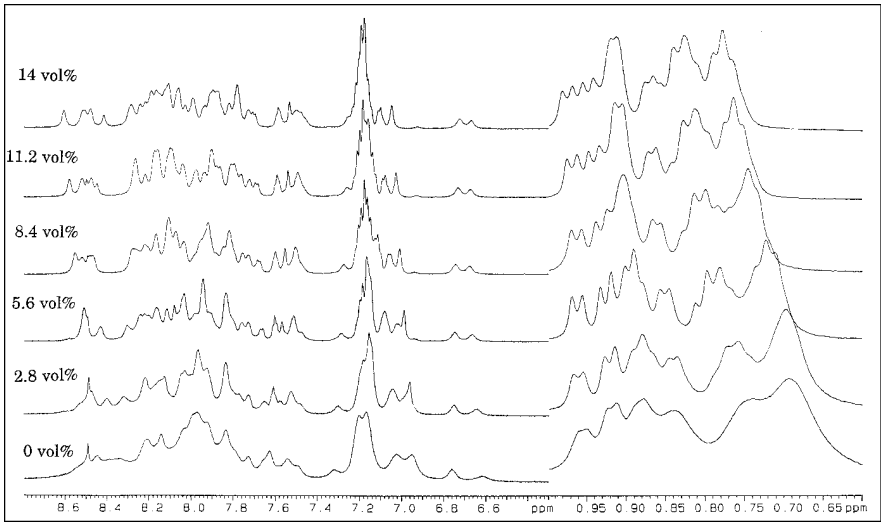
Figure 2



Schematic structure of SA-42 showing the design of the hydrophobic core. NOE contacts between the aromatic protons of Phe35 and Phe38 at the lower end of helix II and the methyl groups of Nle5, Ile9 and Leu12 are evidence for a hairpin conformation. NOE contacts between the aromatic protons of Phe35 and Phe38 and the methyl groups of Nle16 and Nle27 are not possible intramolecularly and provide evidence for antiparallell dimerization [10].

Figure 3

Parts of the ^{1}H NMR spectrum of SA-42 at 323 K and pH 6.3 in aqueous solution and in mixtures of water and TFE. From bottom to top the concentration of TFE is 0, 2.8, 5.6, 8.4, 11.2 and 14 vol% TFE. The methyl region shows decreased chemical shift dispersion and linewidths with increased TFE concentration due to the increased population of non-interacting helices. The amide proton exchange rates are reduced with increasing TFE concentration and the amide resonances are therefore sharpened. Some amide proton resonances are shifted to lower field with increased TFE concentration showing that in aqueous solution the exchange with the solvent is not slow on the NMR timescale. The sharp singlet at 8.48 ppm is due to an impurity.



exchange of the amide protons which severely hampered the possibility of obtaining the necessary information from the ^{1}H NMR spectrum. The ^{1}H NMR spectrum of SA-42 in aqueous solution shows broad lines that become more narrow with increasing temperature (Fig. 4). The polypeptide is therefore in fast exchange between several conformations, which is typical of molten globules.

Circular dichroism and the dimer to monomer transition

The mean residue ellipticity at *222* nm, θ_{222}, is commonly used to assess the helical content of polypeptides and proteins [15–17]. A more negative value of θ_{222} corresponds to a larger helical content. The CD spectrum of SA-42 was recorded as a function of TFE (Fig. 5). The mean residue ellipticity at *222* nm became increasingly negative as a function of TFE and most of the increase took place in the interval from 0 to 10 vol% of TFE. The total increase was less than 10% of the mean residue ellipticity in water. At TFE concentrations larger than 30 vol% no further increase in the negative value of θ_{222} could be seen. The ratio of $\theta_{222}/\theta_{208}$ decreased gradually from 0.95 in water to 0.85 in 70 vol% TFE. Similar changes have previously been interpreted in terms of a transition from a coiled-coil to a single helix [18].

Assignment of the ^{1}H NMR spectrum

The ^{1}H NMR spectrum of SA-42 was assigned in mixtures containing 2.8, 5.6, 8.4, 11.2 and 14 vol% TFE. By design SA-42 contains no repetitive sequences and the assignment was therefore considerably simplified. Typically, all the 36

Figure 4

Parts of the ^{1}H NMR spectrum of SA-42 in aqueous solution at pH 6.3 as a function of temperature. From bottom to top the temperatures are 303, 313, 323, 333, 343 and 353 K. The methyl region chemical shift dispersion and linewidths are decreased with increased temperature due to increased populations of random coils. The amide proton exchange rates are increased with increased temperature and the intensities are reduced due to fast exchange with solvent water. The extreme linewidths at 303 K suggest that SA-42 is close to the coalescence temperature and slow exchange on the NMR timescale.

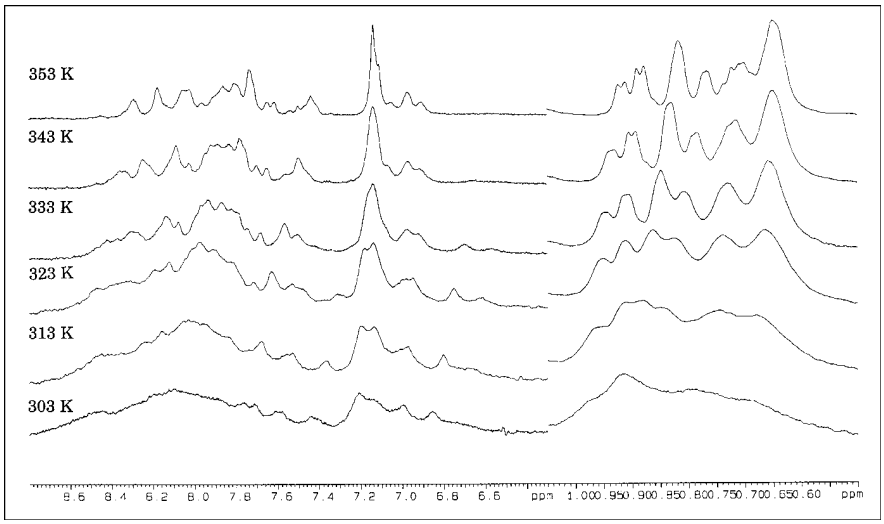
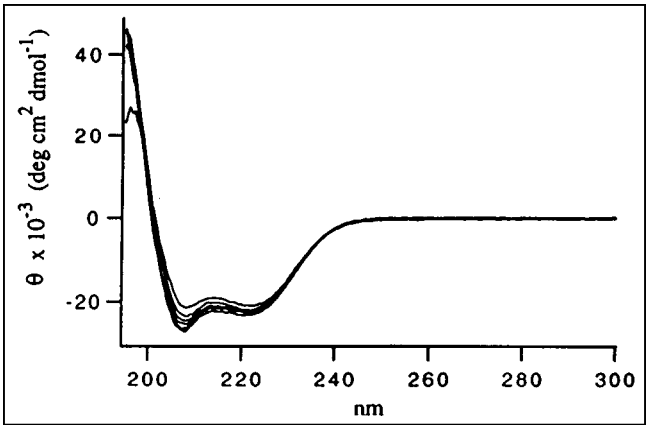
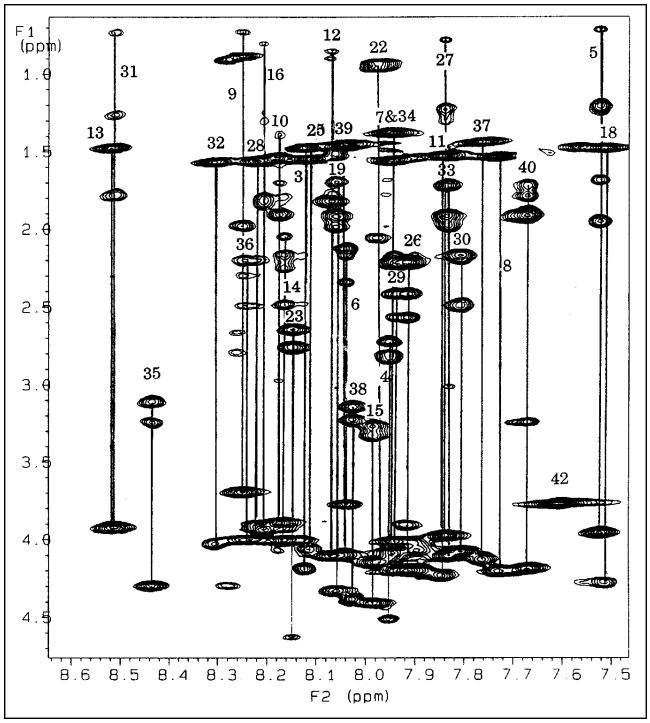


Figure 5



The CD spectrum of SA-42 as a function of TFE concentration at 298 K and pH 6.9. The spectrum in aqueous solution has the least negative mean residue ellipticity at 222 nm and the lowest helical content. The spectra in TFE solutions show increased negative values of the mean residue ellipticity at 222 nm and are closely similar for concentrations of TFE above 30 vol%. From top to bottom the concentrations are 0, 5, 10, 20, 30, 40, 50, 60 and 70 vol% of TFE.

Figure 6



Part of the TOCSY spectrum of SA-42 in 5.6 vol% of TFE in H_2O/D_2O (90/10 v/v) at 323 K and pH 6.3. Each spin system has been labelled with the number in the amino acid sequence (Fig. 1). All spin systems have been assigned from the TOCSY spectrum except for the four α-aminoisobutyric acids which have no α protons. Pro21 and Asn1, which have no amide protons, have been assigned from their α proton chemical shifts. The NH chemical shift of Gly20, 7.92 ppm, was assigned from the NOESY spectrum (Fig. 7). Some amino acid residues close to Pro21 in the amino acid sequence of SA-42 have extra traces with weak intensities due to the *cis*/*trans* equilibrium of proline as reported earlier [10].

spin systems from amino acids with both αH and NH and Pro21 and Asn1 were identified in the TOCSY spectrum (Fig. 6). The NH chemical shifts of α-aminoisobutyric acid (Aib) residues and most of the sequential assignments were obtained directly from the NH–NH region of the NOESY spectrum (Fig. 7). Some readily identifiable amino acids occur only once in the sequence and some appear in less crowded parts of the NMR spectrum. The chemical shifts vary roughly linearly with TFE concentration at different rates for different resonances, and spin systems that overlap in one TFE composition are resolved in others. Approximately 75% of the resonances in aqueous solution could be assigned by extrapolation from the chemical shifts obtained in TFE solution. The main difficulties in assigning the ^{1}H NMR spectrum of a dynamic peptide in aqueous solution arise from fast amide proton exchange rates although conformational equilibria probably also contribute to line broadening. Fast exchange leads to low intensities in the amide regions of e.g. TOCSY and NOESY spectra. This problem was readily resolved by the addition of small amounts of TFE.

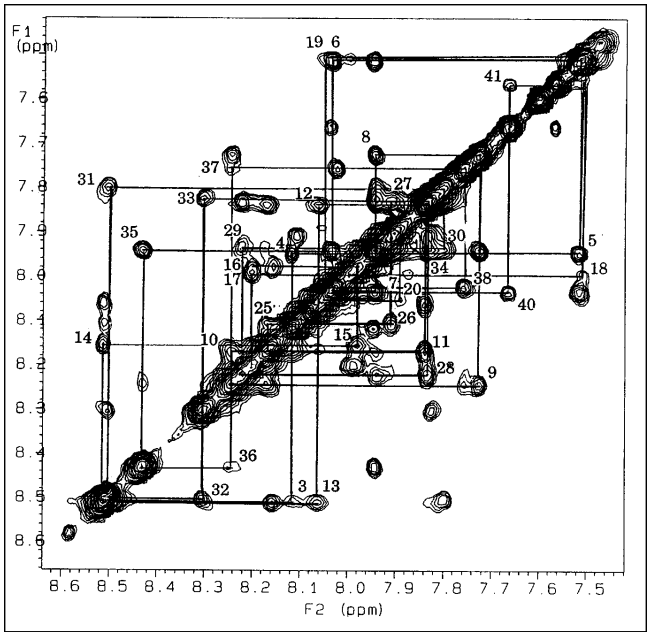
The structure of SA-42

The structure of SA-42 in aqueous solution was previously determined from the observed NOEs and from equilibrium sedimentation ultracentrifugation. It forms a helix-loop-helix hairpin conformation that dimerizes in an antiparallel mode in solution [10]. Medium-range NOEs typical of helix formation (αH NH i,i+3 and i,i+4) [19] were observed from Ala3 to Aib17 and from Ala25 to Arg40. NOEs that connect the phenylalanine aromatic protons and methyl groups of the amino acid residues in the hydrophobic core (Figs 2,8) were observed in aqueous solution and in solvent mixtures with less than 14 vol% of TFE. Observed connectivities between Phe35 aromatic protons and the methyl groups of Ile9 and Leu12 together with the observed NOE contacts between phenylalanine aromatic protons and those of the methyl groups of norleucines Nle16 and Nle27 showed that SA-42 folds into an antiparallel hairpin helix-loop-helix dimer.

The effect of TFE on the supersecondary structure of SA-42

The long-range NOEs have now been studied as a function of TFE concentration (Fig. 8). The intensities of the crosspeaks that connect phenylalanine aromatic protons in helix II with methyl groups of hydrophobic residues in helix I are measures of the lifetime of the hairpin motif. The intensities of the crosspeaks that connect phenylalaine aromatic protons with methyl groups of Nle16 and Nle27 are measures of the lifetime of the dimer. The intensities of all long-range NOEs are gradually decreased

Figure 7



Part of the NOESY spectrum of SA-42 in 5.6 vol% TFE in H_2O/D_2O (90/10 v/v) at 323 K and pH 6.3. The NOESY walk is indicated and the crosspeaks are labelled with the numbers of the amino acid that has the highest number in the sequence of the two that are connected, i.e. the crosspeak labelled 31 connects Gln30 and Leu31. The NH chemical shifts of Gly20 and the α-aminoisobutyric acid residues were assigned from the sequential NOEs.

with increasing TFE concentration and the crosspeaks are too weak to be observed in 30 vol% TFE. The connectivity between Phe35 and Ile9 that provides information about the hairpin conformation is weak but observable in 25 vol% of TFE but the connectivities between phenylalanine and Nle16 and Nle27 are observable only at concentrations of TFE below 15 vol%, due to overlap that arises as the chemical shift dispersion decreases.

At 30 vol% TFE, the mean residue ellipticity of SA-42 is more negative than in water and the helical structure is therefore more developed. The NOE connectivities between the helices are, however, no longer observable and SA-42 therefore exists predominantly as a monomer with two non-interacting helices that are connected by a short loop. These results are in agreement with the conclusion that the change in the ratio of $\theta_{222}/\theta_{208}$ reflects a transition from a tetrahelix bundle to non-interacting helices.

Methyl group chemical shifts as indicators of supersecondary structure

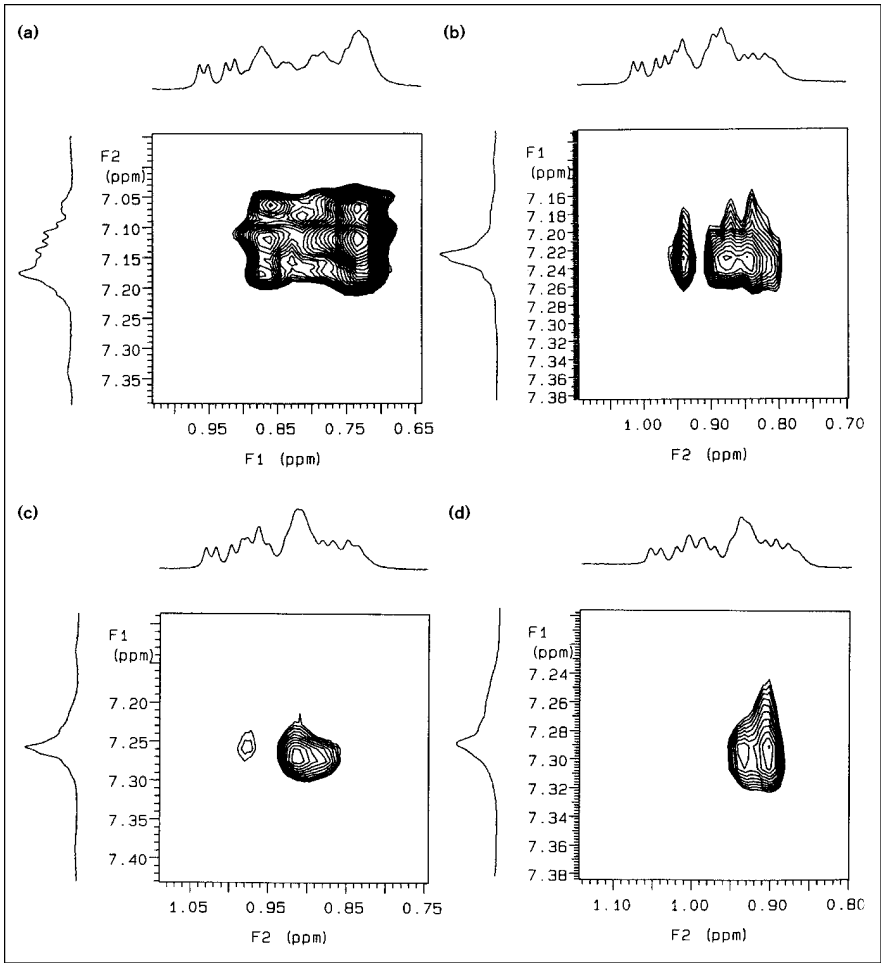
In aqueous solution, in the folded form, the methyl groups are expected to be ordered to some extent whereas at high TFE concentrations SA-42 has little hairpin conformation and the sidechains are rotating freely on the surface of the solvent-exposed helix. The chemical shifts of the methyl groups can therefore be used as indicators of supersecondary structure formation. The chemical shifts of the methyl groups of the amino acid residues designed to make up the hydrophobic core in the folded polypeptide vary with the concentration of TFE. The chemical shift dispersion observed in aqueous solution decreases with the addition of TFE and the chemical shifts in 20 vol% TFE are very close to those of the random coils reported by Wüthrich [19]. Those of the methyl groups of Leu12 are 0.94 and 0.91 ppm which are almost identical to those of a random-coil leucine, 0.94 and 0.90. The chemical shifts of the methyl groups of Leu31, which are close to Phe35 in the helical conformation, are 0.88 and 0.85 ppm, 0.06 ppm from those of Leu12. The methyl groups of Nle5 and Nle27, for which there are no reported corresponding random-coil values, have a chemical shift of 0.89 ppm, which is very close to the expected value of a leucine. The methyl groups of Ile9 at 20 vol% appear at 0.95 and 0.82 ppm, which can be compared to coil shifts of 0.95 and 0.89. Ile9 is the residue that shows crosspeaks to Phe35 at the highest TFE concentration and it may be shifted to some extent even in 20 vol% TFE due to partial formation of supersecondary structure. Val22 methyl groups deviate from coil values by less than 0.02 ppm in all solvents. The similarity of the chemical shifts to those of random coils is surprising since in 20 vol% TFE the amino acid residues are in fact in a helix conformation. It is probably due to the fact that the sidechains are located on the surface of a helix, rotate freely and are exposed to solvent.

At TFE concentrations higher than 20 vol% where there is little hairpin conformation, according to the weakness of observable NOEs, the chemical shifts of all methyl groups shift to higher field with 0.003–0.005 ppm per vol% of TFE due to a common solvent effect. The solvent effect for addition of less than 20 vol% TFE is small. The chemical shifts of the methyl groups of Ile9, Leu12, Leu31, Nle5, Nle16, Nle27 and Val22 relative to random-coil shifts as a function of TFE concentration are presented in Figure 9. The chemical shifts in aqueous solution of the methyl groups of Leu31, the δ-methyl of Ile9 and that of Nle5 deviate the most from the random-coil shifts whereas Nle16 and Nle27 are intermediate and Leu12 appears to be close to the coil value even in aqueous solution.

αH chemical shifts as indicators of secondary structure

The α protons (αH) of the amino acid residues of polypeptides and proteins are commonly used as indices of secondary structure [20]: a negative (upfield) shift relative to that of the random-coil conformation indicates helix formation, whereas a positive (downfield) shift is typical of β-structures or turn formation. The α-proton chemical shift deviation of SA-42 in 5.6 vol% TFE is presented in Figure 10a and indicates helix formation from amino acid residues 4–16 and from 25–40 and a turn/loop from 18–23. The residues 2, 17, 24 and 41 are α-aminoisobutyric acids

Figure 8



The long-range crosspeaks that connect Phe35 and Phe38 aromatic protons with the methyl groups of the amino acid residues in the hydrophobic core of SA-42 recorded in (a) 10, (b) 20, (c) 25 and (d) 30 vol% TFE in H_2O/D_2O (90/10 v/v) at 323 K and pH 6.3.

and have no αH. The effect on helix formation of TFE can thus be monitored from the change in chemical shifts of the α protons as a function of TFE concentration. The difference in αH chemical shifts between 14 vol% TFE and water is small, but larger in helix I than in helix II (Fig. 10b). The repetitive pattern of large and small shifts relative to those of random coils is considered typical of 'bent' helices [21] and is observed for SA-42. The αH chemical shifts do not indicate more pronounced helix formation at the helical ends with increased TFE concentration, but the largest increases in helix formation are detected in the middle segments of each helix (Fig. 10). The amide proton (NH) chemical shifts of the backbone of polypeptides have also been used as indicators of helix distortions [22]. For SA-42 the NH shifts show a repetitive pattern (Fig. 10c) and these shifts also change more in helix I than in helix II as the fraction of TFE is increased (Fig 10d). The NH chemical shifts are therefore compatible with the interpretation that the SA-42 helices are bent although the presented evidence does not provide conclusive evidence for this subtle structural effect.

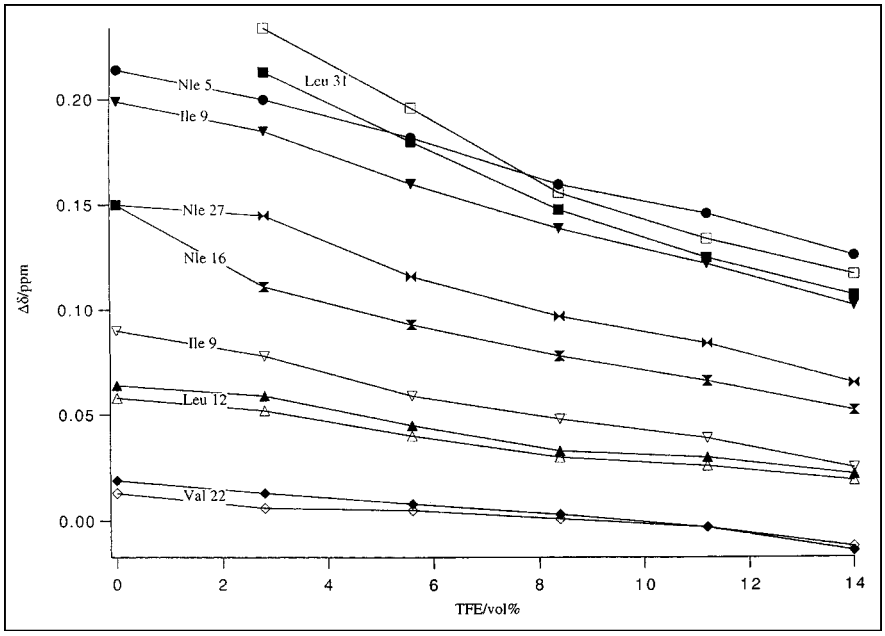
Discussion

The NMR spectroscopic determination of the structure of SA-42 in water and in dilute TFE

In aqueous solution the chemical shift dispersion of SA-42 is significant, but the linewidth and the relaxation properties of the resonances of SA-42 made it very difficult to assign the ^{1}H NMR spectrum. Broadened and overlapping resonances are characteristic of molten globule structures and polypeptides that are approaching native-like properties are the hardest ones to study by NMR, as they are close to coalescence on the NMR timescale. The ^{1}H NMR spectrum of SA-42, which falls in that category, was, however, readily assigned in mixtures of TFE and water between 2.8 and 14 vol% of TFE where the spectral properties were significantly improved (Figs 6,7). TFE titration studies may therefore be of general interest as more designed polypeptides with tertiary structure are engineered.

Due to the assignments in TFE mixtures it was possible to assign a large number of resonances in aqueous solution by extrapolation. But, although more than 75% of the reso-

Figure 9



The variation of methyl group chemical shifts with TFE concentration relative to those of random coils. The justification for comparing methyl group chemical shifts with those of random-coil chemical shifts comes from the observation that in 20 vol% TFE and above, when SA-42 is predominantly folded into two non-interacting helices connected by a loop, the chemical shifts of the methyl groups are very close to those reported by Wüthrich [19] for small peptides without supersecondary structure. The random-coil value of leucine has been used also for norleucine since the chemical shifts of norleucine methyl groups are very similar to those of leucine. The chemical shifts of all three norleucine methyl groups in SA-42 are almost the same in 20 vol% TFE. The deviations from coil shifts suggest that Leu31, Nle5, Nle16, Nle27 and Ile9 are well ordered, whereas Leu12 is not. Val22 methyl group chemical shifts, which do not deviate from coil shifts by more than 0.02 ppm, show that Val22 is not a part of the hydrophobic core. There are no NOE contacts observed for Val22 methyl groups.

nances in aqueous solution were assigned, it was difficult to resolve the NOE crosspeaks since they were severely broadened and of weak intensity. However, in solutions containing 2.8 vol%, 0.7 mol%, or more of TFE the appearance of the NOESY spectrum was much improved and the secondary and supersecondary structures of SA-42 were determined from NOEs in all solvent mixtures used as discussed above. The difference between the structure determined in 2.8 vol% TFE and that in water then had to be estimated, and this was done through an analysis of the magnitude of the chemical shift changes. We are interested in the structure determined in 2.8 vol% because it is the one that is the closest to that in water.

The structure of SA-42 in dilute TFE is similar to the one in water

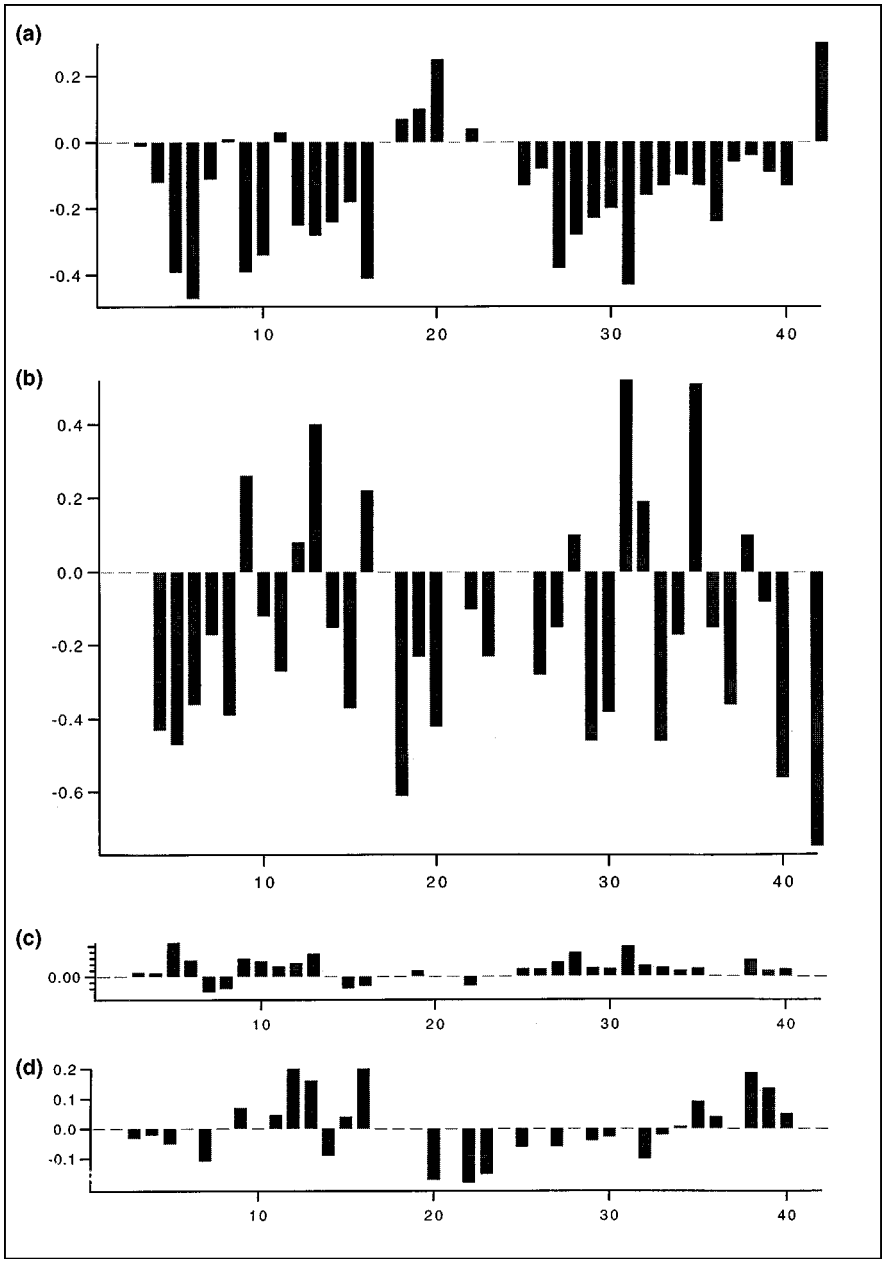
The chemical shift of nuclei in fast exchange is the weighted average of the shifts of the exchanging conformers and a chemical shift change shows that the relative concentrations of the exchanging conformers have changed. The chemical shifts of the methyl groups of the norleucines, isoleucine and leucines of SA-42 were compared with those expected for random coils (Fig. 9) so that the degradation of the supersecondary structure by TFE could be monitored. In aqueous solution the deviations from coil shifts varied between 0.02 ppm (Val22) and 0.30 ppm (Leu31) (Fig. 9). In 20 vol% TFE the chemical shifts approached those of random coils.

The difference between the chemical shifts in 20 vol% TFE and the ones in aqueous solution are due to the transition from four-helix bundle to non-interacting helices. The difference between the chemical shifts in 2.8 vol% TFE and the ones in aqueous solution amounts to 10–20% of the difference that corresponds to the complete transition. The difference between the structure of SA-42 in 2.8 vol% and that in water is therefore small. The NOEs that are used to determine the structure of SA-42 in water are also present in dilute TFE solution. Structural information obtained for SA-42 in dilute TFE is therefore relevant to the structure in water.

Probing the packing of the hydrophobic core

The deviations from the coil shifts were also used as an indication of how well the residues were ordered in the hydrophobic core of the four-helix bundle. Leu31, Nle5 and the δ-methyl group of Ile9 were found to show the largest deviations and they are all situated at the lower end of helix I and II in close proximity to the phenylalanines in the folded peptide. The resonances of the methyl groups of Leu12 are close to their coil shifts although Leu12 can be as close in space as Ile9 and Nle5 to the aromatic residues and it is suggested that the packing of Leu12 is less well ordered. The methyl groups of Nle16 and Nle27 are not as much shifted as that of Nle5 and the proximity of Leu12, Nle16 and Nle27 suggests that the region close to the loop including Leu12 is less well organized than the rest of the hydrophobic core. Proximity to the aromatic residues is not enough to induce large shifts. In 20 vol% TFE, where there is little supersecondary structure in SA-42, the chemical shifts of the methyl groups of Leu31 are

Figure 10



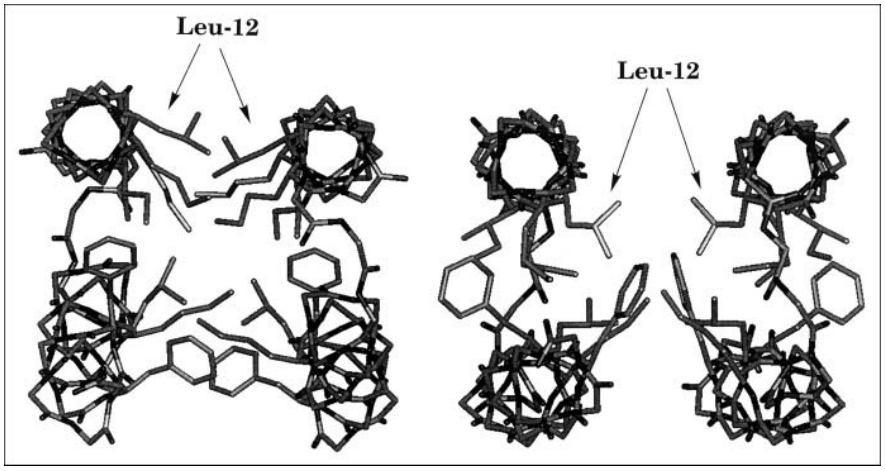
Deviation from random-coil values of (a) αH chemical shifts and (b) NH chemical shifts of the amino acid residues of SA-42 in 5.6 vol% TFE in H_2O/D_2O (90/10 v/v) at 323 K and pH 6.3. The overall change in chemical shifts in the TFE concentration range 0–14 vol% is given for the (c) αH and (d) NH chemical shifts.

close to their random-coil shifts although Leu31 is close to Phe35 in space. A deviation from the coil shift is therefore indicative of an environment that is at least partially ordered. The ring current effects of His15 may of course also contribute in the dimer to the observed chemical shift of Leu31, but only if a relatively well defined structural geometry exists. The observed NOE contacts show that in the predominant conformations in the four-helix bundle all of the methyl groups, except for those of Val22, are close to the aromatic rings of Phe35 and Phe38.

Some support for the interpretation that the sidechain of Leu12 is less well ordered than those of the other hydrophobic residues comes from inspection of modelled dimer structures of SA-42 (Fig. 11). Two extreme dimer conformations may be modelled. The hydrophobic residues of helix I may be oriented intramolecularly towards those of helix II or intermolecularly towards those of helix I, and a range of intermediate conformations are also possible. Leu12 interacts less well with the other hydrophobic residues and appears to be somewhat exposed to solvent. Chemical shift information may in this

Figure 11

Modelled dimer structures of SA-42 showing possible orientations of helices in the four-helix bundle. The labelled sidechain is that of Leu12. The modelling of the dimer was not carried out using bent or coiled helices.



respect perhaps prove useful in the discrimination between possible folds and provide information that can be used in the development of molten globules. More work is needed, however, to establish the potential usefulness of this chemical shift index.

Helix–helix interactions

The magnitude of the variation of αH chemical shifts with the change in structure of SA-42 is small, less than 0.1 ppm, suggesting that there is little distortion of the helices in the formation of the helical interfaces. The CD spectrum changes somewhat due to increased helicity as SA-42 changes from a four-helix bundle to a monomer with non-interacting helices. The increase in the mean residue ellipticity is, however, less than 10% with an addition of 10 vol% TFE. The ratio of $\theta_{222}/\theta_{208}$ for SA-42 changes from 0.95 to 0.85 in the interval 0–70 vol% TFE, which is a smaller change than that reported as typical for the transition from a coiled coil to a single helix. Typical values are 1 and 0.8 [18]. These results are in agreement with the conclusion that the variation in helical structure is small with changes in supersecondary structure and that the secondary structure is already well developed in water. There is thus little evidence to suggest that helix I and helix II have adapted much in the folding process to achieve optimum complementarity.

On the present level of modelling it is an ambitious task to try to engineer perfect complementarity between helical hydrophobic interfaces. It may therefore be advantageous to design four-helix bundles with less than maximum helix propensity if the goal is a well defined tertiary structure so that each helix has more potential for adaptation. The helix propensity of SA-42 is much larger than that of α_2D [9] according to several algorithms for helix prediction [23,24]. Whether there is a correlation between helix propensity and actual 'stiffness' of the helices remains to be experimentally verified in more cases.

Conclusions

The use of small amounts of TFE improves the quality of NMR spectra of molten globules. A large amount of information becomes available once the NMR spectrum has been assigned, including the secondary and supersecondary structures. SA-42 is by all standards a molten globule. However, the temperature dependence of the ^{1}H NMR spectrum suggests that it is very close to coalescence and slow exchange on the NMR timescale, which is a very good criterion for being 'better' than a molten globule. The concept of molten globules is wide and includes a wide range of polypeptides and proteins. Some are far from being well packed and some probably need very little in terms of structural improvement to develop into native-like proteins. For the 'good' molten globules it is necessary that structural information in solution is available so that the reasons why the polypeptide does not fold like a protein can be identified and remedied. It is possible that the use of TFE can provide the necessary improvement of NMR spectra and open the way towards obtaining structural information for a large number of designed polypeptides which can then be developed into native-like proteins.

Materials and methods

Synthesis of SA-42

The synthesis and purification of SA-42 has been described in detail previously [10]. It was synthesized on a Biosearch 9600 automated peptide synthesizer using t-BOC protected amino acids and a Pam-linked resin, and it was cleaved in anhydrous liquid HF on a Teflon vacuum line. It was purified by size-exclusion chromatography and reversed-phase and ion-exchange HPLC and identified by electrospray mass spectrometry and amino acid analysis.

Circular dichroism

CD spectra were recorded on a Jasco J-720 spectropolarimeter which was routinely calibrated with d-campher sulfonic acid [25]. Samples were prepared by dilution, with water and TFE, of an aqueous stock solution of 800 μM peptide in water, pH 6.95. Peptide stock solutions were made up by weight, assuming a water content of 25% in the lyophilized peptide, as seen from earlier quantitative amino acid analyses, followed by addition of the solvent with a volumetric pipette. The concentration was then corrected by using the known mean residue ellipticity of SA-42 [10] in aqueous solution at 293 K, –25 000 deg cm^2 $dmol^{-1}$. Spectra were recorded in a 1 mm cell using a peptide concentration of 80 μM. The solvent compositions used were 0, 5, 10, 20, 30, 40, 50, 60 and 70 vol% TFE in water. Spectra of four scans were collected at room temperature and baseline corrected. The mean residue ellipticity of SA-42 is independent of concentration in aqueous solution [10].

NMR spectroscopy

NMR spectra were recorded on a Varian Unity 500 NMR spectrometer equipped with a matrix shim system MHU-303 from Resonance Research Inc. The one-dimensional (1D) ^{1}H NMR spectrum and the TOCSY and NOESY spectra in mixtures of TFE and H_2O/D_2O (90/10 v/v) (0–30 vol% TFE) were recorded with water suppression accomplished by transmitter presaturation for 1.5 s. Acquisition times of 1, 0.2 and 0.2 s, respectively, a sweep width of 6500 Hz and a 90° pulse of 10.4 μs were used. The spin-lock pulse in the 'clean TOCSY' experiments was 12.5 μs and the window function was 25 μs. The spin lock time was 80 ms. In the NOESY experiments, the mixing time was 250 ms. The TOCSY and NOESY experiments were carried out in the phase-sensitive mode according to States, Habercorn and Ruben. Data matrices of 2×350 increments were collected with 16 transients in each increment. The peptide concentration was 2 mM prior to TFE addition. 1D TOCSY and NOESY ^{1}H NMR spectra were recorded for every solvent composition investigated.

NMR spectra were recorded as described above at 323 K in 0.0, 2.8, 5.6, 8.4, 11.2 and 14 vol% TFE-d_3 in H_2O/D_2O (90/10 v/v) pH 6.3 after TFE addition, and 10, 20, 25 and 30 vol% TFE-d_3 in H_2O/D_2O (90/10 v/v) pH 3.4 prior to TFE addition. TFE affects the pH of the solution probably due to a solvent effect on the pKas of the ionizable residues of SA-42. The effect is small, however, amounting to a few tenths of a pH unit and was not corrected for. Interexperimental errors due to these corrections would probably have been larger than those that arise from the small variations in pH. The CD spectrum of SA-42 is independent of pH in the interval from 4.4 to 8 and the effects outside this interval are small. The effect of pH on the ^{1}H NMR spectrum is clearly observable but negligible for the changes in pH induced by TFE. An NOE build-up curve with mixing times of 50, 100, 150, 200 and 250 ms was recorded for the solution containing 14 vol% TFE. No signs of spin diffusion were detected. ^{1}H NMR spectra of SA-42 in aqueous solution at pH 5.75 were recorded at 303, 313, 323, 333, 343 and 353 K.

Acknowledgements

We are indebted to Per Ahlberg for stimulating discussions. Financial support from the Swedish Natural Science Research Council is gratefully acknowledged.

References

1. Bryson, J.W., *et al.*, & DeGrado, W.F. (1995). Protein design: a hierarchic approach. *Science* 270, 935–941.
2. Mutter, M. & Vuilleumier, S. (1989). A chemical approach to protein design – template assembled synthetic proteins (TASP). *Angew. Chem. Int. Ed.* 28, 535–554.
3. Johnsson, K., Allemann, R.K., Widmer, H. & Benner, S.A. (1993). Synthesis, structure and activity of artificial, rationally designed catalytic polypeptides. *Nature* 365, 530–532.
4. Baltzer, L., Lundh, A.-C., Broo, K., Olofsson, S. & Ahlberg, P. (1996). Polypeptides with supersecondary structure as templates in rational catalyst design. Catalysis of self functionalization by designed helix-loop-helix motifs. *J. Chem. Soc. Perkin Trans.* 2, 1671–1676.
5. Kaumaya, P.T.P., Berndt, K.D., Heidorn, D.B., Trewhella, J., Kezdy, F.J. & Goldberg, E. (1990). Synthesis and biophysical characterization of engineered topographic immunogenic determinants with αα topology. *Biochemistry* 29, 13–23.
6. DeGrado, W.F. & Lear, J.D. (1990). Conformationally constrained α-helical peptide models for protein ion channels. *Biopolymers* 29, 205–213.
7. Åkerfeldt, K.S., Kim, R.M., Camac, D., Groves, J.T., Lear, J.D. & DeGrado, W.F. (1992). Tetraphilin: a four-helix proton channel built on a tetraphenylporphyrin framework. *J. Am. Chem. Soc.* 114, 9656–9657.
8. O'Neil, K.T., Hoess, R.H. & Degrado, W.F. (1990). Design of DNA-binding peptides based on the leucine zipper motif. *Science* 249, 774–778.
9. Raleigh, D.P., Betz, S.F. & DeGrado, W.F. (1995). A *de novo* designed protein mimics the native state of natural proteins. *J. Am. Chem. Soc.* 117, 7558–7559.
10. Olofsson, S., Johansson, G. & Baltzer, L. (1995). Design, synthesis and solution structure of a helix-loop-helix dimer – a template for the rational design of catalytically active polypeptides. *J. Chem. Soc. Perkin Trans.* 2, 2047–2056.
11. Struthers, M.D., Cheng, R.P. & Imperiali, B. (1996). Design of a monomeric 23-residue polypeptide with defined tertiary structure. *Science* 271, 342–345.
12. Llinas, M. & Klein, M.P. (1975). Charge relay at the peptide bond. A proton magnetic resonance study of solvation effects on the amide electron density distribution. *J. Am. Chem. Soc.* 97, 4731–4737.
13. Jasanoff, A. & Fersht, A.R. (1994). Quantitative determination of helical propensities from trifluoroethanol titration curves. *Biochemistry* 33, 2129–2135.
14. Buck, M., Radford, R.S. & Dobson, C.M. (1993). Partially folded state of hen egg white lysozyme in trifluoroethanol: structural characterization and implications for protein folding. *Biochemistry* 32, 669–678.
15. Adler, A.J., Greenfield, N.J. & Fasman, G.D. (1973). Circular dichroism and optical rotatory dispersion of proteins and polypeptides. *Methods Enzymol.* 27, 675–711.
16. Chen, Y.-H, Yang, J.T. & Chau, K.H. (1974). Determination of the helix and β form of proteins in aqueous solution by circular dichroism. *Biochemistry* 13, 3350–3359.
17. Engel, M., Williams, R.W. & Erickson, B.W. (1991). Designed coiled-coil proteins: synthesis and spectroscopy of two 78-residue α-helical dimers. *Biochemistry* 30, 3161–3169.
18. Lau, S.Y.M., Taneja, A.K. & Hodges, R.S. (1984). Synthesis of a model protein of defined secondary and quaternary structure. *J. Biol. Chem.* 259, 13253–13261.
19. Wüthrich, K. (1985). *NMR of Proteins and Nucleic Acids.* Wiley, New York.
20. Wishart, D.S., Sykes, B.D. & Richards, F.M. (1991). Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.* 222, 311–333.
21. Blanco, F.J., *et al.*, & Nieto, J.L. (1992). NMR chemical shifts: a tool to characterize distortions of peptide and protein helices. *J. Am. Chem. Soc.* 114, 9676–9677.
22. Zhou, N.E., Zhu, B.-Y., Sykes, B.D. & Hodges, R.S. (1992). Relationship between amide proton chemical shifts and hydrogen bonding in amphipathic α-helical peptides. *J. Am. Chem. Soc.* 114, 4320–4326.
23. Richardson, J.S. & Richardson, D.C. (1989). *Prediction of Protein Structure and the Principles of Protein Conformation.* (Fasman, G.D., ed), Plenum Press, New York.
24. Munoz, V. & Serrano, L. (1995). Elucidating the folding problem of helical peptides using empirical parameters. II. Helix macro dipole effects and rational modification of the helical content of natural peptides. *J. Mol. Biol.* 245, 275–296.
25. Johnson, W.C., Jr. (1990). Protein secondary structure and circular dichroism: a practical guide. *Proteins* 7, 205–214.